Analyzing Strawberry Pricing Strategies Across Canadian Vendors*

Price Variations Driven by Supplier Strategies and Seasonal Demand in Canadian Grocery Markets

Tianning He

November 25, 2024

This study investigates the pricing of strawberries across eight major suppliers in Canada: Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart, and Save-On-Foods. Using a dataset containing variables such as current price, old price, and month, a Bayesian regression model was applied to analyze the relationships between pricing strategies, seasonal trends, and vendor-specific practices. Results reveal significant price variation among suppliers, with old price showing a strong positive correlation with current price, reflecting the impact of prior pricing strategies on current prices. Seasonal trends captured by the month variable indicate subtle but consistent fluctuations in pricing, highlighting the influence of supply and demand dynamics. These findings provide actionable insights into supplier-specific pricing strategies and market behavior, offering valuable guidance to consumers and stakeholders in the Canadian strawberry market.

1 Introduction

Strawberries are one of the most important small berries in Canada and the United States(Macoun and McCubbin (1919)), and pricing of strawberries in the Canadian retail market varies widely due to factors such as supplier strategies, promotional practices, and temporal trends. This investigation examines the determinants of strawberry prices, focusing on differences between different suppliers, the role of discounts, and seasonal effects. By utilizing a Bayesian linear regression model, the study examines the relationship between key predictors (supplier, original price, promotion, and time) and the current price of strawberries. The Bayesian approach not only provides parameter estimates but also quantifies uncertainty,

*Code and data are available at: https://github.com/HeTianning/The-price-of-strawberries

thus providing a nuanced understanding of price dynamics in competitive and seasonal markets.

The main objective of this analysis is to estimate and explain the factors that influence retail strawberry prices in Canada. The dependent variable current price reflects the Canadian dollar price per pound of strawberries. The key predictors include supplier identity (categorical), which captures supplier-specific pricing strategies; old price (continuous), which represents the original price before promotion, and month (continuous), which captures temporal or seasonal trends(Hodgdon et al. (2024)). By modeling these relationships, the study aims to quantify how these variables affect pricing and explore their combined impact on strawberry price changes.

The results show significant differences in supplier-specific pricing strategies, with Metro offering higher prices than other suppliers, consistent with its premium market positioning. Time trends indicate a gradual but consistent increase in strawberry prices over time, possibly reflecting seasonal demand or inflationary pressures. Furthermore, the analysis shows a strong relationship between original and discounted prices, suggesting that promotional pricing generally preserves relative price hierarchies among suppliers. Bayesian models confirm strong convergence and robust posterior estimates, ensuring the reliability of these findings.

Understanding the factors driving strawberry prices is critical for both consumers and retailers. For suppliers, insights into pricing dynamics can inform strategies to optimize revenue while remaining competitive. For consumers, this analysis provides valuable information to navigate the market, highlighting opportunities to save costs by selecting suppliers or purchasing during promotions. More broadly, these findings contribute to the study of the fresh produce market, where pricing is influenced by a complex interaction between supplier strategies, seasonal supply, and consumer behavior. This study lays the foundation for future research that could incorporate additional factors, such as geographic differences or organic certification, to further enrich our understanding of pricing dynamics in Canadian retail.

The rest of the paper is structured as follows. Section **??** describes the data collection process, including data sources, sampling strategy, and preprocessing steps taken to ensure data quality. Section **??** outlines the methodological framework, focusing on the Bayesian linear regression model, its formulation, and diagnostic checks. Section **??** presents the main findings of the analysis, focusing on specific supplier pricing strategies, seasonal trends, and the impact of sales and promotions on strawberry prices. Section **??** discusses the implications of these findings, placing them in the broader context of consumer behavior and supplier strategies in the fresh produce market.

2 Data

To conduct the analysis, `R`, a versatile statistical programming language, was utilized for simulating, testing, downloading, cleaning, and modeling data (R Core Team 2023). Several

powerful libraries supported various stages of the analysis. The `tidyverse` package suite (Wickham et al. 2019), which includes `dplyr` for data manipulation and `ggplot2` for visualization, was central to handling and exploring the dataset. Data cleaning and structuring were enhanced using `janitor`, while `arrow` facilitated efficient handling of large datasets by enabling seamless reading and writing of data (Apache Arrow 2021). The `here` package streamlined file management, ensuring robust and reproducible workflows (Müller 2020).

Visualization and reporting were further supported by `knitr` (Xie 2021), which enabled dynamic report generation, and `modelsummary` (Arel-Bundock 2022), which provided professional-quality summaries for statistical models. Bayesian modeling and advanced statistical analysis were conducted using `rstanarm` (Goodrich et al. 2022), while posterior diagnostics and visualizations were enhanced with `bayesplot` (Gelman, Gabry, et al. 2021). Testing and quality assurance of code were carried out with `testthat` (Wickham 2011), ensuring the reliability and reproducibility of results.

By integrating these libraries the analysis leveraged a comprehensive ecosystem of tools, demonstrating the strength and flexibility of R in handling complex data-driven research workflows.

2.1 Overview

The data for this analysis was provided by Project Hammer (Filipp 2024), which was collected through a process of scraping website UI screens from vendors within the Toronto community. It contains data on product details (e.g., product name, brand, current and old price, and unit of product) for 8 different grocery stores (Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart Canada, and Save-On-Foods). Project Hammer's data is available from February 28, 2024, and since the data is collected by extracting information from grocery store websites, there may be missing data, as discussed in appendix. Through the method of screen scraping, the raw data is intended to measure and reflect grocery store vendor decisions and product pricing trends.

For this analysis, the eight vendors - Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart Canada, and Save-On-Foods - and their pricing for strawberries were considered. The variable "price" for each supplier in this analysis reflects the current price of their product at the time the data was downloaded.

2.2 Measurement

The dataset provides a detailed and structured framework for analyzing strawberry pricing dynamics in the Canadian market, offering critical insights into supplier-specific pricing strategies and seasonal trends. The central variable of interest is the current price, a continuous measure representing the price of strawberries in Canadian dollars at the time of observation.

This variable acts as the dependent variable in the analysis and reflects real-time vendor pricing decisions, including adjustments for promotions, sales, or other market-driven factors. The current price is essential for understanding how vendors position their products in response to competition and consumer demand.

Complementing this, the dataset includes the old price, another continuous variable that captures the price of strawberries before any discounts or promotional offers. The old price serves as a benchmark for evaluating vendor markdowns, sales strategies, and pricing flexibility. By calculating the difference between the old price and the current price, the analysis can quantify the extent and frequency of price reductions, revealing patterns in promotional behavior. For instance, larger discounts may indicate aggressive marketing tactics during periods of high supply or low demand. Together, current and old prices allow for a nuanced exploration of how pricing evolves over time and in response to market conditions.

The dataset also incorporates vendor as a categorical variable, identifying one of eight major suppliers: Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart, and Save-On-Foods. This variable is pivotal for inter-vendor comparisons, enabling the analysis to uncover unique pricing strategies, competitive behaviors, and market segmentation. By examining differences in pricing across these suppliers, the analysis can identify which vendors consistently offer lower prices, adopt aggressive discounting practices, or align their pricing with premium branding.

Another critical variable is month, recorded as a numeric measure that captures the time of observation on a monthly scale. This variable facilitates the investigation of seasonal trends, allowing for the analysis of how strawberry prices fluctuate throughout the year. Seasonal dynamics may reflect changes in supply (e.g., peak harvest periods), demand (e.g., holidays or special occasions), and vendor-specific pricing adjustments. By integrating month with pricing data, the analysis can identify whether prices tend to peak or dip during specific times of the year and how these patterns vary among vendors.

These variables collectively provide a rich dataset for conducting both granular and aggregate analyses of strawberry pricing. The inclusion of current and old prices enables an in-depth exploration of vendor-specific pricing adjustments, while the categorical vendor variable allows for competitive benchmarking across suppliers. The month variable introduces a temporal dimension, facilitating the study of seasonal effects on pricing. Together, this dataset offers a comprehensive foundation for understanding the complex interplay between market forces, supplier behavior, and seasonal trends, yielding actionable insights for consumers seeking the best value and suppliers aiming to optimize their pricing strategies.

2.3 Outcome variable

The primary outcome variable in this study is current price, which represents the current price of strawberries in Canadian dollars. This continuous variable serves as the dependent variable and captures the prices consumers pay at the point of purchase, reflecting vendor-specific pricing strategies and promotional adjustments. Additionally, old price is included as a secondary

outcome variable, representing the original price before any discounts. The relationship between these two variables highlights the frequency and magnitude of price reductions, offering insights into vendor markdown strategies.

The analysis focuses on understanding how current price varies with factors such as vendor and month. The categorical variable vendor identifies the supplier and captures differences in pricing strategies across eight vendors, ranging from premium to budget-focused retailers. These variations help reveal how suppliers target different market segments through pricing. The numeric variable month enables the exploration of seasonal trends and temporal fluctuations, reflecting how supply and demand dynamics influence pricing throughout the year.

By examining both current price and old price, the study provides a detailed view of promotional activity and vendor behavior. A significant gap between the two prices indicates aggressive promotional strategies, while minimal differences suggest stable pricing policies. Additionally, incorporating the month variable allows for insights into seasonal peaks or dips in strawberry prices, such as higher prices during off-seasons or lower prices during harvest months. Together, these variables offer a comprehensive understanding of strawberry pricing dynamics in Canada's retail market, helping inform consumer choices and vendor pricing strategies.

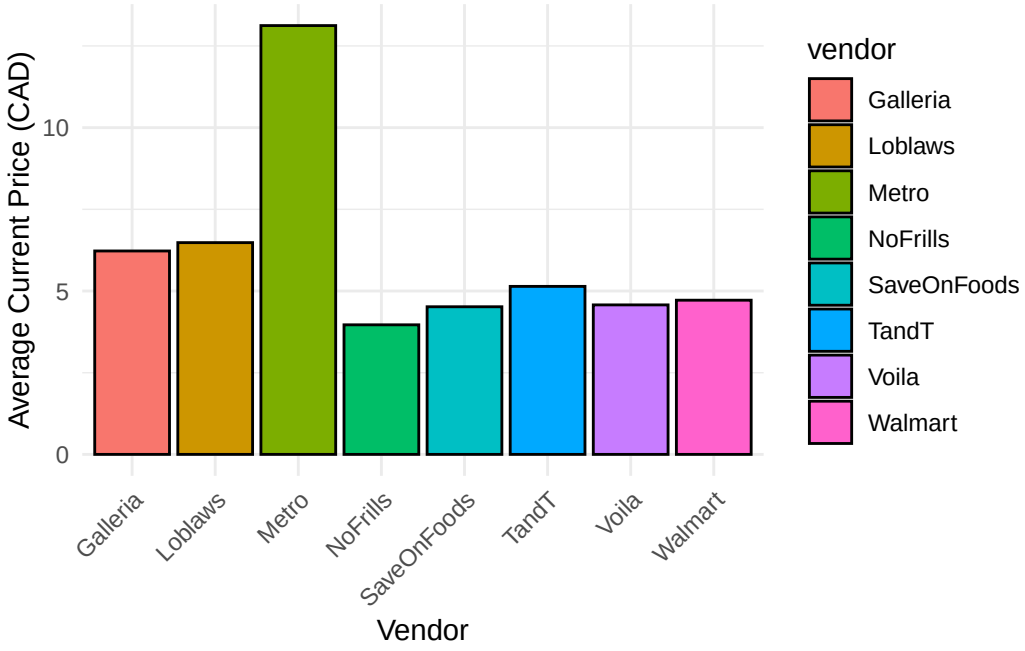

Figure 1: Average Current Price by Vendor

The graph(Figure **??**) illustrates the average current price of strawberries across various vendors, highlighting significant price variations. `Metro` stands out with the highest average price, indicating a focus on premium offerings or higher-quality products, while `No Frills`

and `Save-On-Foods` have the lowest prices, appealing to cost-conscious shoppers. Vendors like `Loblaws`, `T&T`, and `Voila` fall into the mid-range pricing category, suggesting a balance between affordability and quality. These differences reflect distinct market strategies, with premium vendors targeting quality-focused consumers and budget-friendly vendors catering to price-sensitive customers. The data provides valuable insights for consumers to make informed purchasing decisions and for vendors to refine their pricing strategies to better compete in their respective market segments.

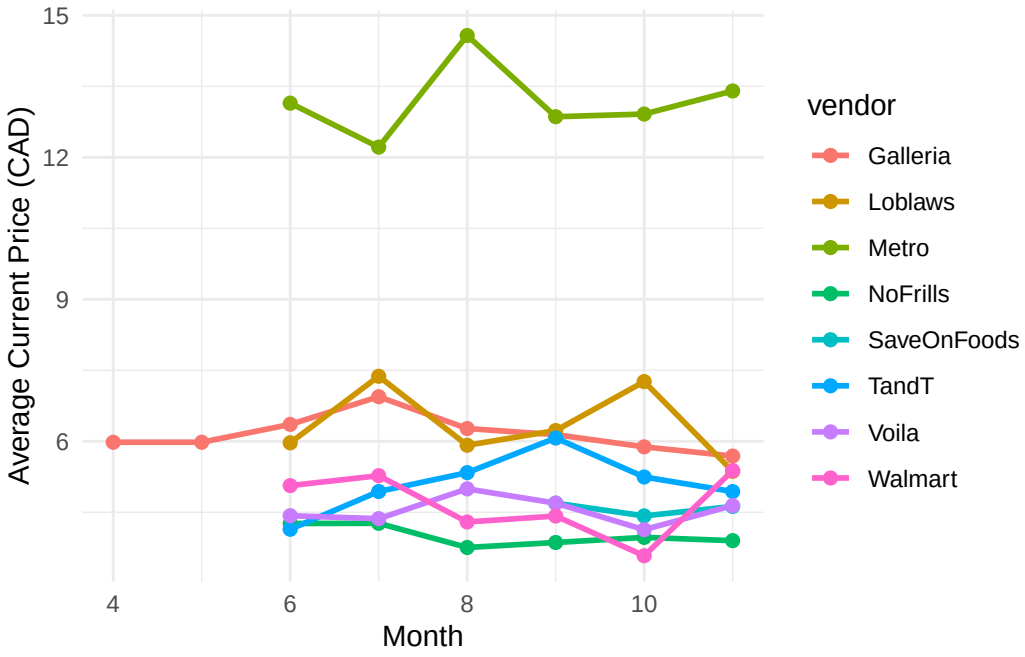

Figure 2: Monthly Trend of Current Prices by Vendor

This graph(Figure **??**) shows the monthly trend of average current strawberry prices across different vendors over several months. `Metro` consistently exhibits the highest prices, with noticeable fluctuations, peaking in mid-year, indicating potential seasonal effects or changes in supply and demand. In contrast, budget-friendly vendors like `No Frills` and `Save-On-Foods` maintain the lowest prices, demonstrating pricing stability aimed at cost-conscious consumers. Mid-range vendors such as `Loblaws`, `T&T`, and `Voila` show relatively stable but slightly variable trends, with occasional peaks suggesting targeted pricing adjustments. Overall, the graph highlights clear segmentation among vendors, with premium vendors like `Metro` adapting dynamically to market conditions, while budget vendors maintain consistent affordability. This trend could reflect differing business strategies to cater to diverse consumer preferences.

2.4 Predictor variables

Predictor Variables: Exploring Factors Influencing Strawberry Prices

1. Vendor

The vendor variable is categorical, identifying the supplier of strawberries. Each vendor employs unique pricing strategies, leading to significant variability in pricing across the eight vendors (Voila, T&T, Loblaws, No Frills, Metro, Galleria, Walmart, and Save-On-Foods). Premium vendors like Metro and Galleria tend to target quality-conscious consumers, reflected in their higher average prices. In contrast, budget-oriented vendors like No Frills and Save-On-Foods maintain lower prices to attract cost-sensitive shoppers. The variability in pricing, shown through boxplots, highlights how vendor identity influences price distributions and captures the strategic segmentation of the Canadian strawberry market.

2. Old and Current Prices

The dataset includes two continuous price variables: old price (original price) and current price (price after any promotions or discounts). The difference between these two variables provides insights into promotional behavior and markdown strategies. Vendors like No Frills and Save-On-Foods typically show small gaps between old price and current price, suggesting limited use of deep discounts, whereas Metro and Walmart exhibit larger gaps, reflecting more aggressive promotional activity. Boxplots comparing old price and current price by vendor demonstrate how discounts influence consumer pricing perceptions and vendor competition.

3. Month (month)

The month variable is numeric, capturing the month of observation. Seasonal patterns play a crucial role in strawberry pricing, with supply and demand fluctuations driving price variability. For instance, prices may drop during peak harvest months and increase during off-season periods when strawberries are scarce. Line charts tracking average monthly prices show clear temporal trends, with peaks and dips aligning with these seasonal dynamics. By analyzing month in combination with other predictors, this variable helps uncover how vendors adjust pricing strategies to reflect seasonal demand and supply conditions.

The interplay of vendor, old price, current price, and month allows for a more comprehensive analysis of pricing strategies. For example, vendors like Metro may use larger discounts (old price vs. current price) during peak seasons to maintain competitiveness, while budget vendors like No Frills may maintain consistent pricing throughout the year. A combined analysis using bar charts and line plots highlights how vendor strategies vary with time, providing insights into the dynamic nature of strawberry pricing.

The inclusion of variables like vendor, old price, current price, and month enriches the analysis by offering multiple perspectives on pricing strategies. Vendor captures the segmentation of suppliers into premium and budget categories, reflecting strategic decisions about market

positioning. The dual price variables (old price and current price) allow for a deeper understanding of promotional activity and highlight the extent to which vendors rely on discounts to influence consumer behavior.

The addition of month introduces a temporal dimension, enabling the analysis of seasonal trends that are common in agricultural markets. Combining these predictors provides a holistic view of pricing dynamics, revealing how vendors adapt to seasonal supply fluctuations, consumer demand, and competitive pressures.

By leveraging these predictor variables, the study offers actionable insights into how strawberry prices are shaped by market forces, vendor decisions, and seasonal trends. These findings are crucial for both consumers seeking to optimize purchases and vendors aiming to refine their pricing strategies to remain competitive.

3 Model

3.1 Purpose of the Model

The purpose of this model is to identify and quantify the factors influencing the current price of strawberries across various vendors in Canada. The analysis evaluates how key predictors, such as vendor, original price, month, and their interactions, affect pricing. Specifically, the model aims to:

- Quantify the impact of each predictor (e.g., vendor, original price, and month) on strawberry prices.
- Analyze how the gap between original (`old_price`) and discounted (`current_price`) prices varies across vendors.
- Examine seasonal trends in strawberry prices to identify periods of high or low prices.
- Provide actionable insights into vendor pricing strategies and market dynamics.

3.2 Model Description

The chosen model is a **Bayesian linear regression model**, which builds on the simplicity of traditional linear regression while incorporating the Bayesian framework. This allows for probabilistic insights and robust uncertainty quantification, making it well-suited for analyzing the factors driving strawberry prices. Further background details and diagnostics are included in Appendix **??**.

Predictors in the Model

- **Vendor (`vendor`)**: A categorical variable identifying the supplier of strawberries. Different vendors adopt distinct pricing strategies that influence strawberry prices.